DOCUMENT RESUME

ED 398 262 TM 025 216

AUTHOR Cross, Lawrence H.; Frary, Robert B.
TITLE Hodgepodge Grading: Endorsed by Students and Teachers Alike.
PUB DATE Apr 96
NOTE 15p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New York, NY, April 9-11, 1996).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Educational Assessment; *Grades (Scholastic); *Grading; High Schools; *High School Students; Junior High Schools; *Junior High School Students; *Secondary School Teachers; Student Evaluation; Surveys; *Teacher Attitudes; Teacher Education; Theory Practice Relationship

ABSTRACT

Previous research clearly documents that teachers often award what S. M. Brookhart (1991) has referred to as a "hodgepodge grade of attitude, effort, and achievement." This paper reports on a survey of grading practices involving 310 middle and high school teachers of academic subjects in a single school system. Also surveyed were 7,367 middle and high school students from the same system. The results largely validate the findings of earlier studies. Substantial majorities of the teachers reported "hodgepodge" grading practices. More important, the students largely confirmed and supported the hodgepodge grading practices reported by their teachers. These results are contrasted with grading practices widely recommended in measurement texts followed by a discussion of how measurement specialists may be missing the mark in their efforts to communicate their views to teachers, school administrators, and the general public. (Contains 4 tables and 11 references.) (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Hodgepodge Grading; Endorsed by Students and Teachers Alike

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Lawrence H. Cross[1] and Robert B. Frary
Virginia Polytechnic Institute and State University

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
LAWRENCE H. CROSS
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

ED 398 262

Presented at the annual meeting of the
National Council on Measurement in Education
New York, April, 1996

Abstract

Previous research clearly documents that teachers often award what Brookhart (1991) has referred to as a "hodgepodge grade of attitude, effort and achievement" (p. 36). This paper reports on a survey of grading practices involving 310 middle and high teachers of academic subjects in a single school system. Also surveyed were 7367 middle and high school students from the same system. The results largely validate the findings of earlier studies. Substantial majorities of the teachers reported "hodgepodge" grading practices. More important, the students largely confirmed and supported the hodgepodge grading practices reported by their teachers. These results are contrasted with grading practices widely recommended in measurement texts followed by a discussion of how measurement specialists may be missing the mark in their efforts to communicate their views to teachers, school administrators, and the general public.

[1]Educational Leadership and Policy Studies, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061-0302. Tel. 540-231-9725. E-mail CROSS@VTVM1.CC.VT.EDU

TM025216

BEST COPY AVAILABLE

Hodgepodge Grading;
Endorsed by Students and Teachers Alike

Lawrence H. Cross and Robert B. Frary
Virginia Polytechnic Institute and State University

School marks and grading have been the source of continuous controversy since the turn of the century (Cureton 1971, Ebel & Frisbie 1986, Hopkins, Stanley, & Hopkins, 1990). Ebel and Frisbie identified three factors contributing to the controversies surrounding marks and grading: the technical difficulties of measuring educational achievement, differences in educational philosophies, and the conflict in roles arising when teachers serve as both advocates and judges. Despite the controversies or the reasons for them, there is widespread agreement among measurement specialists that grades, at least in academic subjects, should be based exclusively on measures of current achievement and that growth, ability, effort, conduct, and other non-achievement factors should not be considered.

Though measurement specialists may agree, recent research reviewed by Brookhart (1994) clearly documents the prevalence of grading practices that diverge from this recommendation. In an attempt to explain this phenomenon, Brookhart (1991) suggested that teachers may use nonachievement factors in determining grades in order to mitigate negative social consequences associated with what she considers to be inappropriate use of grades, such as determining eligibility for nonacademic privileges either at home or at school. She suggests that the solution is not a "*hodgepodge* grade of attitude, effort, and achievement" (p. 36, italics added) but lies in encouraging teachers to employ valid grading practices and in admonishing parents, administrators and the general public about improper uses of grades. While we support the spirit of her recommendations, we see little reason to believe that the measurement community will have much impact on the misuse of grades, when we apparently have had so little effect on teacher grading practices. Indeed, Brookhart later acknowledged that "more training, by itself, will not cause grading practices to conform completely to recommendations" (1994, p. 290).

We would like to suggest that hodgepodge grading may serve to protect not only students, but also teachers, from negative professional or social consequences. Incorporation of factors such as effort or ability may prevent what might be viewed as excessive numbers of failing grades or grades that might appear to be biased with respect to some ethnic or social group. This supposition aside, it is undoubtedly true that many teachers blatantly use grades based on factors such as conduct, attitude, and even attendance to control student behavior.

That grades at the elementary school level are likely to reflect a hodgepodge of both academic and social achievement factors was noted more than 35 years ago by the eminent sociologist Talcott Parsons (1959). In describing the nature of school achievement expected of elementary school children, he identified two major components that he labeled "cognitive" and "moral." He noted that:

> . . . in the elementary grades these two primary components are not clearly differentiated from each other. Rather, the pupil is evaluated in diffusely general terms; a *good* student is defined in terms of a fusion of the cognitive and the moral components, in which varying weights are given to one or the other. Broadly speaking, then, we may say that the "high achievers" of the elementary school are both the "bright" pupils, who catch on more easily to their more strictly intellectual tasks, and the more "responsible" pupils, who, "behave well" and on whom the teacher can "count" in her difficult problem of managing her class. (p. 304)

Though Parsons does not specifically address grading at the secondary level, it seems reasonable to assume that at this level he would have expected greater importance to be assigned to the "cognitive" rather than the "moral" component.

No distinction between elementary and secondary schools is made by Airasian and Jones (1993) who call on the measurement community to broaden the perspective of what constitutes classroom measurement and assessment and to be more sensitive to the realities of the classroom. They suggest that "context independent, best measurement practices . . . are generally dismissed by teachers as 'unrealistic,' 'impractical,' or more bluntly, 'not relevant to classroom needs'" (p. 241). They note that it is difficult for most teachers, especially elementary teachers, ". . . to separate their knowledge and perceptions of students from their grading judgments, and so many do not" (p. 250). These observations no doubt hold true for a large majority of teachers across all grade levels. Moreover, as Parsons suggests, it may be natural to assume that both the "moral" and the "cognitive" factors are considered at the elementary level, at which other means of communicating academic achievement per se are available.

Contrary to the general theme of Airasian and Jones (1993), we are not prepared to advocate equal billing for the "moral" and "cognitive" components in grading at the secondary level, especially among teachers of academic subjects. We believe the measurement community has an obligation help these teachers appreciate the need to make a clear distinction between measured academic achievement and their perceptions of the "whole child." The latter may be crucial for the many instructional decisions a teachers must make, as noted by Airasian and Jones, but it is ill-advised, in our opinion, to give credence to the "moral" component when determining grades in academic subjects at the secondary level. Because of the importance placed on these grades, either for educational or occupational decisions, grades should communicate as objectively as possible the levels of educational attainment in the subject. To encourage anything less, in our opinion, is to distort the meaning of grades as measures of academic achievement, at a time when the need for clarity of meaning is greatest.

Method

The research reported here was conducted at the invitation of a school district that was concerned with various grading issues at the secondary level, including the use of hodgepodge grading practices. The district, with a population of approximately 150,000, incorporates a large military base and is a vacation destination. A unique aspect of this study is that students, as well as teachers participated in the survey. Although Brookhart (1994) states that "confounding of effort and achievement is prevalent in practice, and strongly endorsed by teachers and students" (p. 290), only two of the studies cited in her review involved students. One involved only 13

secondary students, and the other involved only elementary students. By contrast, in the present study, all middle and high school teachers and students in the district were requested to participate, and very large majorities of both groups did. Thus it was possible to compare student and teacher perceptions of grading practices in a single school system. Of special interest was the extent to which students and teachers share common understandings of grading practices.

Two instruments were developed, one for teachers and one for students. The teacher survey was modeled after one used for a survey of Virginia secondary teachers of academic subjects (Frary, Cross, & Weber, 1993). This survey was the only one reviewed by Brookhart (1994) that was based on a large-scale random sample. Hence, one goal of the present study was to compare results with the statewide results from Virginia. As in the Virginia study, the teachers were asked to describe their actual grading practices, their opinions regarding testing and grading, and to identify their school level, teaching experience, and their teaching field. The instrument contained 54 forced-choice items.

The student survey, also in forced-choice format, inquired as to the importance that their teachers placed on various grading factors, as well as their satisfaction with the grading process. It contained 51 items, including 14 identical to those on the teacher survey. The students were also asked to provide various personal information, such as gender, ethnicity, parents' education, perceived level of academic achievement, and effort expended on studying, homework, etc.

The surveys were distributed in the spring of 1994 to all middle and high school students and teachers. Responses were obtained from 465 teachers, including 226 middle school teachers and 239 high school teachers. Table 1 shows a breakdown of the "major teaching area" for these teachers. However, in what follows, findings are reported only with respect to teachers of academic subjects (152 middle school teachers and 158 high school teachers). Responses from teachers of non-academic subjects were not considered in subsequent analyses because: a) the consequences of grades awarded by these teachers no doubt differ from the consequences associated with grades awarded by academic teachers, b) the grading practices of these teachers might be expected (and were found) to differ from those of teachers of academic subjects, and c) we wished to compare responses of teachers from this survey to responses obtained from academic teachers in our statewide survey of teachers in Virginia.

Similarly, only responses from high school students in the college preparatory curriculum were included in the analyses reported here (N=3161). This decision reflected our belief that grade transcripts play a much more important role in the opportunities available to college bound students than for students not planning to go to college. Moreover, the perceptions of teacher grading practices by college bound students would focus more directly on academic courses than would the perceptions of non-college bound students. It might be noted that the student survey asked about grading practices of teachers in general and was not subject specific. All middle school respondents were included (N=4122) since tracking had not yet formally occurred for these students.

In addition to response frequency distributions, matrices of interitem correlations were produced for the student and faculty responses, using dummy variables as appropriate for items with nonordinal response options. For example, the correlations between dummy variables representing the major teaching areas and all other items were studied to evaluate response

variation possibly associated with teaching area. The tables that follow show the response percentages to selected survey items across all teachers and students from the district surveyed. For comparison purposes, the response percentages for the same or similar items from our survey of Virginia teachers are also included in the tables. The respondents from the current survey are referred to as "district teachers" or "district students" so as to maintain the anonymity of the (non-Virginia) school system surveyed. Crosstabulations were created to illustrate correlational findings of interest and are discussed in the text.

The findings are presented and discussed with respect to what we will refer to as recommended practices. These are practices represented by substantial consensus among measurement textbook authors such as Ebel and Frisbie (1986); Hopkins, Stanley, and Hopkins (1990); and Gronlund (1985).

Results

The extent to which teachers and students endorsed hodgepodge grading can be seen by inspecting the response percentages reported in Table 2.[2] Listed in this table are responses to items that assess the extent to which extraneous factors are considered when determining grades. Whereas "recommended practice" would urge teachers to ignore ability when determining grades, only 49% of the district teachers agreed that this is the ideal. Moreover, 72% of the district teachers indicated that they raised the grade of low-ability students. The only exception to this finding was that mathematics teachers were less likely to consider ability (54% v. 72% overall).

Curiously, when responses to these two items, which asked about actual versus ideal practice, were crosstabulated, there were substantial percentages of teachers whose practice was inconsistent with their response concerning the ideal. Specifically, 36% of the 86 teachers who indicated that they do *not* consider ability indicated that ideally ability should be considered. Conversely, of the 217 teachers who indicated that they *do* consider ability, 49% indicated that ideally ability should not be considered.

The students also seemed to be confused over the role ability should play in determining grades. A majority (55%) of the students embraced the notion that it is fair for a teacher to take a student's ability into consideration in determining a report card grade. Yet, 62% indicated agreement with the statement above that ideally ability should not be considered.

As noted by Hopkins et al. (1990), "achievement in relationship to aptitude is an untenable basis for marking, despite the obvious appeal the ideas has for many educators" (p.329). If nothing else, the above findings attest to the "obvious appeal" to which Hopkins et al. refer.

Recommended practice would urge teachers not to consider growth or improvement when determining grades. Not only are growth measures notoriously unreliable (Ebel, 1979; Gronlund, 1985; Hopkins et al., 1990), but their use can lead to motivational problems as

[2]In Table 2 and those that follow, some of the survey item stems have been paraphrased and some response options have been collapsed to facilitate communication of outcomes. Copies of the actual instruments are available upon request to the first author.

discussed by Ebel and Frisbie (1986). Only 20% of the district teachers indicated they would not consider growth (independent of final achievement level), and another 20% indicated that they consider growth to a substantial extent.

Recommended practice asks that teachers ignore effort when determining grades. Nevertheless, 25% of the district teachers indicated that they raise report grades for high effort "fairly often." In contrast, relatively few teachers lower grades for lack of effort among high ability students. The responses of the students tend to corroborate these findings in that 33% indicated that most of their teachers consider effort. Our view of this matter is that the level of effort expended should be directly reflected in the level of achievement attained and there is no need to consider effort beyond this effect.

If grades are to represent current levels of achievement, then poor conduct and attitudes ought to be dealt with separately and not be considered when determining grades. Yet, 39% of the district teachers acknowledge taking conduct/attitude into consideration when determining report card grades. A surprisingly large percentage of students (71%) endorsed this practice and a majority of the students indicated that some (30%) or most (33%) of their teachers did so. However, as noted by Ebel (1979), unless ". . . improving students' attitudes toward something, or improving their willingness to put forth effort is one of the specific objectives of the course . . ., attitude and effort probably should be excluded from consideration in determining the mark to be assigned" (p.238).

In contrast to the grading practices reported above, when asked whether achievement, effort and conduct should be reported separately, 81% of the teachers and 70% of the students agreed or tended to agree with this assertion. Though Brookhart (1994) points out that the use of separate grades does not constitute a panacea and may introduce problems of a different sort (p. 297), we note that no recent research has addressed the efficacy of this practice.

Table 3 contains response percentages to items concerning homework. Homework was reported to have a strong influence on grades by 27% of the district teachers and to have a moderate influence by another 46%. An exception to these findings was the foreign language teachers, who were less likely to assign homework (25% indicated giving no homework) and were less likely to grade homework (35% versus 52% overall). A majority (52%) of the district teachers indicated that their homework assignments, are graded either by the class (16%) or by themselves (36%), but the rest either check homework only for completion (43%) or not at all (5%). These findings are in contrast to recommended practice. As Hopkins et al. (1990) suggests, "ordinarily, homework should not have much influence on students' marks, unless they do not do the assignments" (p. 333). Moreover, he notes that "Students should be able to get help on homework assignments, but such help also means that the resulting product is not a good indicator of the student's mastery (p. 333). By contrast, only 40% of the district teachers agreed that "grading homework is unfair, because the teacher doesn't know the extent to which the student may have gotten help in preparing it."

Though most teachers encourage class participation, it is generally not recommended practice to consider class participation when determining grades due to the subjective nature of the assessment. A possible exception to this admonition among teachers of academic subjects is foreign language instruction, for which class participation may play a central role in developing

oral skills in the language. Nonetheless, as shown in Table 3, class participation was rated by a majority of district teachers as having a moderate (39%) or strong (14%) influence on grades. Forty-six percent reported that they relied on their recollection and adjusted only some grades, and 41% indicated that they recorded participation and adjusted some (20%) or all grades (21%). A majority of the students (63%) recognized that some (27%) or most (36%) of their teachers count class participation and most were "very" (33%) or "somewhat satisfied" (35%) with the way their teachers grade class participation.

Table 4 contains response percentages to items that assess perceptions regarding tests and test scores. The first set of items focuses on the meaning of percentage scores. It is our contention that, for most teacher-made tests, percentage scores provide no more than a basis for ranking students (Frary, et al., 1993). Few teacher-made tests are capable of yielding criterion-referenced interpretations (even if designed to do so) and, and in our opinion, it is erroneous to interpret such scores as representing the percentage of knowledge a student possesses with regard to: a) a specified domain of knowledge, b) content taught, or even c) content covered by the test. At one level, most of the district teachers seem to understand this. A majority (63%) of these teachers endorsed the statement that "percentage grading scales are usually arbitrary, since 90% correct on an easy test may represent performance comparable to 80% on a more difficult test of the same topic." An even larger majority (77%) of district teachers endorsed the statement that a teacher can make grades come out as they like by adjusting the difficulty of the test questions. Comparable percentages of students endorsed both of these statements. In contrast to the understanding implied by the above responses, 85% of the district teachers endorsed the statement that "The percentage of correct answers on a classroom test is usually a good estimate of the percentage of the material covered by the test that a student has learned." Surely they must realize that the difficulty of the test questions will influence a student's percent-correct score quite independently of how much the student has learned.

Responses of the Virginia teachers were at almost exactly the same level as the district teachers with respect to the arbitrariness of percentage grading scales. However, the Virginia teachers were less supportive of the notion that percentage scores represent the percentage of material learned (59% vs. 85%). The issue of teacher adjustment of test difficulty to obtain desired grades was handled differently on the Virginia survey in that it was tied to the efficacy of district-wide percentage grading scales. Though 77% of the district teachers supported the idea that teachers could adjust test difficulty to obtain desired grades, only 40% of Virginia teachers did so when this action was proposed as a reason for ineffectiveness of district-wide percentage grading scales. (On another survey item, Virginia teachers had expressed widespread support for district-wide percentage grading scales.)

As Hopkins et al. (1990) observe, "A rigid adherence to the conventional percentage could discourage teachers from including many items from the higher taxonomy levels . . ." (p. 323). Indeed, teachers tend to include many easy items so as to insure that the percent correct scores will range between 60% or 70% to 100%, not appreciating the need to include difficult items in their tests. More difficult tests will rank order the scores more reliably than easy tests and provide better discrimination across grade categories. In contrast to the Virginia teachers, a majority of the district teachers seem to appreciate the need for difficult tests, in that only 33% endorsed the statement that "An easier test with scores in the 70% to 100% range will rank students more accurately in the subject tested than a difficult test with scores in the range of 35%

to 90%." However, this response-level difference only reflect the difference in the wording of this item for the two surveys. This suspicion is given credence by the fact that only 40% of the district teachers seemed to appreciate the need for a few test questions sufficiently difficult that less able students will be unable to answer no matter how hard they have studied. No doubt, many found this notion to be cruel and heartless, even though it represents sound psychometric advice. But, again, wording differences for the item may have been partially responsible for the response percentages.

It was interesting to note that student responses concerning grading practices and level of satisfaction with them did not vary meaningfully with respect to most of the variables reflecting personal characteristics of students. That is, weaker and stronger students, students belonging different ethnic groups, male and female students, etc., gave highly similar survey responses. However, middle school students tended to endorse use of extraneous factors more so than high school students.

Discussion

This study is important for two reasons beyond the local uses made of the survey. First, it provides an extensive validation of the findings of Frary, Cross, and Weber (1993) with regard to hodgepodge grading. Second, and perhaps more important, it shows that students appeared to understand and endorse these practices as reasonable.

If teachers were to embrace grading practices as recommended by measurement specialists, surely more valid indicators of academic achievement would result. However, there appears to be considerable resistance to doing so, even among those who have been trained in recommended practices Brookhart (1994). The current survey attests to the limited impact training is likely to have. One of the district teacher survey items asked "How much instruction have you received in classroom testing and grading, either formal coursework or inservice training?" Responses to this item were, for all practical purposes, unrelated to responses to the other items on the survey instrument. Thus, the 32% who indicated having had "Moderate to large amount" of instruction in these topics were just as inclined to embrace hodgepodge grading as the 9% indicated "None," or the 52% who indicated "A small amount."

Recognizing that instruction in recommended practices alone may not be the answer, Brookhart (1994) notes that "Present recommendations for grading do not take into account the teacher's need to manage classrooms and motivate students" (p. 299). She suggests several lines of research that may inform us about the dynamic between teaching, learning, and grading that may lead to more realistic recommendations. However, we are not convinced that the "limited perspective" (Brookhart, 1994, p. 290) represented by recommended practices should be abandoned in favor of practices that accommodate the nuances of each classroom and each pupil, and certainly not in secondary academic classrooms.

These points aside, we must ask, if hodgepodge grading is so deplorable, why haven't students, parents, administrators, or the general public called for reform? It may well be that they share a common understanding that grades often do, in fact, represent a hodgepodge of attitude, effort, conduct, growth, and achievement, and that is what they expect and endorse. It seems likely that they expect uniformly high performance in these areas to result in high grades but understand

and accept that poor performance in any one area may result in a somewhat lower grade. That grades are likely to be biased by the subjectivity teachers use in assessing and combining these ingredients may be perceived as less of a concern than to discount effort, ability, attitudes, conduct, and growth as irrelevant considerations. Although the measurement community has an obligation to warn of the technical complexities of measuring and including these factors in grades, we have to recognize that to students, teachers, administrators, and parents there is considerable face validity to grades which include extraneous factors, despite the great potential for subjective teacher bias to distort the meaning of grades.

Given these circumstances, measurement specialists might well attempt communicating with a broader audience concerning the technical aspects of grading. It may well be that many parents, administrators, and even teachers have never seriously thought about the negative aspects of hodgepodge grading. Though measurement specialists discuss these issues in technical journals, others offer advice that is often at odds with recommended practice in journals widely read by teachers and administrators. For example, Guskey (1994) urges teachers to base their grading on three learning criteria: product (what students know and are able to do), process (effort or work habits, class participation, attendance), and progress (improvement or learning gain). Although he dutifully notes that "most researchers and measurement specialists ... recommend using product criteria exclusively" (p. 17), the reader is clearly left with the impression that it is simply a matter of personal preference as to which criteria should be used. Indeed, with advice like this, how can we hope to overcome hodgepodge grading without reaching a wider audience?

BEST COPY AVAILABLE

References

Airasian, P. W. & Jones, A. M. (1993). The teacher as applied measurer: Realities of classroom measurement and assessment. *Applied Measurement in Education, 6*, 241-254.

Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice, 10*(1), 35-36.

Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7*, 279-301.

Cureton, L. W. (1971). The history of grading practices. *Measurement in Education*, 2(4), 1-8.

Ebel, R. L. (1979). *Essentials of Educational Measurement* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall, Inc.

Ebel, R. L., & Frisbie, D. A. (1986). *Essentials of Educational Measurement* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Frary, R. B., Cross, L. H. & Weber, L. J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: implications for instruction in measurement. *Educational Measurement: Issues and Practice, 12*(3), 23-30.

Gronlund, N. E. (1985). *Measurement and evaluation in teaching*. New York: Macmillan.

Guskey, T. R. (1994). Making the grade: What benefits students? *Educational Leadership, 52*(2), 15-20.

Hopkins, K. D., Stanley, J. C. & Hopkins, B. R. (1990). *Educational and psychological measurement and evaluation* (7th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Parsons, T. (1959). The school class as a social system: Some of its functions in American society. *Harvard Educational Review, 29*, 297-318.

Table 1 - Teaching Assignments of District Respondents

	Number of Teachers					
		School Level		College Preparatory Courses Taught		
Major Teaching Area	Total	Middle	High	Most/All	About Half	Few/None
English	91	49	42	31	11	48
Mathematics	85	48	37	22	8	53
Science	57	22	35	19	10	27
Social Studies	56	24	32	19	4	31
Foreign Language	21	9	12	10	2	8
Vocational	65	20	45	8	2	52
Arts	22	14	8	3	0	19
Health/Phys. Ed.	22	11	11	5	2	14
Other	46	29	17	6	6	37
Total	465	226	239	123	39	289

Table 2 - Extraneous Grading Factors

Factor	Sample	Survey Item (Paraphrased)	Responses
Ability	Virginia teachers	Student ability should be considered.	66% agree or tend to agree.
	District teachers	Do you evaluate achievement in terms of ability?	72% yes
		Ideally, it should not be necessary to consider ability in determining a grade.	49% agree or tend to agree.
	District students	Fair for teachers to consider ability in assigning grades?	55% yes
		Ideally, it should not be necessary to consider ability in determining a grade.	62% agree or tend to agree.
Growth	Virginia teachers	Amount a student gains should be considered.	85% agree or tend to agree.
	District teachers	Do you consider growth or improvement?	20% no, 60% a little, 20% a lot
Effort	Virginia teachers	Exceptionally high or low effort should be recognized in final grade.	66% agree or tend to agree.
	District teachers	Do you raise report card grades for high effort?	17% no, 58% rarely, 25% often
		Do you lower report card grades for low effort?	89% no, 7% only for high ability students, 4% yes in all cases
	District students	How satisfied are you with teachers' use of effort to determine grades?	63% very or somewhat satisfied
		How many of your teachers consider effort in determining grades?	12% none, 28% some, 33% most, 27% DK
Behavior/ Conduct	Virginia teachers	Laudatory or disruptive behavior should be considered in final grades.	31% agree or tend to agree
	District teachers	Do you consider conduct or attitude in grades?	39% yes
	District students	How many of your teachers take conduct or attitude into consideration for grading?	15% none, 30% some, 33% most, 25% DK
		Do you think conduct or attitude should be taken into consideration for grading?	37% yes, good and bad 28% good only, 6% bad only, 29% no
Separate Grades	District teachers	Report cards should report achievement effort, and conduct separately.	81% agree or tend to agree
	District students	Report cards should report achievement effort, and conduct separately.	70% agree or tend to agree

Note. Virginia teachers N = 536
District teachers (academic major teaching areas only) N = 310
District students (middle school and high school college preparatory only) N = 7367

Table 3 - Nontest Grading Factors

Factor	Sample	Survey Item (Paraphrased)	Responses
Homework	Virginia teachers	Influence of daily homework on grades.	24% little/none, 71% important, 5% most influential
	District teachers	How much influence does homework have on grades.	27% little/none, 46% moderate, 27% strong
		How do you determine homework grades?	5% not given/graded, 43% completion only, 16% graded by class, 36% teacher graded
		Grading homework is unfair due to help that may have been received.	40% agree or tend to agree
	District students	How satisfied are you with the way teachers assign an grade homework?	80% very or somewhat satisfied
		How much homework, project/term paper help do you get from your family?	18% none, 39% little, 43% fair amount/a lot
Class Participation	Virginia teachers	Influence of class participation on grades.	44% little/none, 51% important, 5% most influential
	District teachers	How much influence does class participation have on grades?	46% little/none, 39% moderate, 14% strong
		How do you use class participation in determining grades?	14% not used, 46% recollection only, 41% record and average in
	District students	How many of your teachers count class participation in grading?	36% most, 27% some, 11% few/none, 26% DK
		How satisfied are you with the way your teachers grade class participation?	68% very or somewhat satisfied

Note. Virginia teachers N = 536
District teachers (academic major teaching areas only) N = 310
District students (middle school and high school college preparatory only) N = 7367

Table 4 - Score Interpretation and Test Difficulty Perceptions

Factor	Sample	Survey Item (Paraphrased)	Percent Agree or Tend to Agree
Score interpretation	Virginia teachers	Percentage grading scales meaningless--unclear what 90% is a percent of.	63%
		Test scores indicate percentage of topic covered by test that students have learned.	59%
		District-wide percentage grading scales ineffective because teachers can adjust test difficulty to get desired grades.	40%
	District teachers	Percentage grading scales arbitrary--90% on easy test may be equal to 80% on a harder test	63%
		Teacher can make grades come out as desired by adjusting test difficulty.	77%
		Percent correct on test is good estimate of percentage of material covered by test that student has learned.	85%
	District students	Percentage grading scales arbitrary--90% on easier test may be equal to 80% on a harder test	70%
		A teacher can make grades come out as desired by adjusting test difficulty.	64%
Effects of test difficulty level	Virginia teachers	Letter grades can be assigned more confidently with score range 30%-90% rather than 60%-100%.	33%
		To challenge better students it is desirable to ask test questions some students may be unable to answer no matter how much they study.	28%
		A harder test (mean=60%) is fairer to students than an easier one (mean=80%).	24%
	District teachers	A test with a score range of 70%-100% will rank students more accurately than one with a score range of 35%-90%.	33%
		To challenge better students it is desirable to ask test questions some students may be unable to answer no matter how much they study.	40%

Note. Virginia teachers N = 536
District teachers (academic major teaching areas only) N = 310
District students (middle school and high school college preparatory only) N = 7367